Osteoporosis International (2019) 30:1433–1443
https://doi.org/10.1007/s00198-019-04935-6

ORIGINAL ARTICLE

Validation of distal radius failure load predictions by homogenized- and micro-finite element analyses based on second-generation high-resolution peripheral quantitative CT images

A. J. Arias-Moreno[1,2] · H. S. Hosseini[3] · M. Bevers[1] · K. Ito[1] · P. Zysset[3] · B. van Rietbergen[1]

Received: 10 September 2018 / Accepted: 5 March 2019 / Published online: 17 April 2019

Abstract

Summary This study developed a well-standardized and reproducible approach for micro-finite element (mFE) and homogenized-FE (hFE) analyses that can accurately predict the distal radius failure load using either mFE or hFE models when using the approaches and parameters developed in this study.

Introduction Micro-FE analyses based on high-resolution peripheral quantitative CT (HR-pQCT) images are frequently used to predict distal radius failure load. With the introduction of a second-generation HR-pQCT device, however, the default modelling approach no longer provides accurate results. The aim of this study was to develop a well-standardized and reproducible approach for mFE and hFE analyses that can provide precise and accurate results for distal radius failure load predictions based on second-generation HR-pQCT images.

Methods Second-generation HR-pQCT was used to scan the distal 20-mm section of 22 cadaver radii. The sections were excised and mechanically tested afterwards. For these sections, mFE and hFE models were made that were used to identify required material parameters by comparing predicted and measured results. Using these parameters, the models were cropped to represent the 10-mm region recommended for clinical studies to test their performance for failure load prediction.

Results After identification of material parameters, the measured failure load of the 20-mm segments was in good agreement with the results of mFE models ($R^2 = 0.969$, slope = 1.035) and hFE models ($R^2 = 0.966$, slope = 0.890). When the models were restricted to the clinical region, mFE still accurately predicted the measured failure load ($R^2 = 0.955$, slope = 1.021), while hFE predictions were precise but tended to overpredict the failure load ($R^2 = 0.952$, slope = 0.780).

Conclusions It was concluded that it is possible to accurately predict the distal radius failure load using either mFE or hFE models when using the approaches and parameters developed in this study.

Keywords Bone strength · Distal radius · Finite element analysis · HR-pQCT · Micro-FE · Osteoporosis

Introduction

Micro-finite element (FE) analyses based on high-resolution peripheral quantitative CT (HR-pQCT) images are nowadays commonly performed in clinical studies to measure bone mechanical properties in vivo (for an overview, see [1, 2]). With this approach, an approximately 1-cm region of the distal radius or distal tibia is imaged at 82 μm voxel size and a Laplace-Hamming filtering followed by a thresholding procedure is used to generate a 3D reconstruction of the bone tissue [3]. This reconstruction is then converted to a micro-FE model from which stiffness, failure load, and load transfer parameters are calculated for compressive loading conditions [4]. In most studies, linear elastic analyses are performed, which cannot simulate the actual failure process during the compression.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00198-019-04935-6) contains supplementary material, which is available to authorized users.

✉ B. van Rietbergen
B.v.Rietbergen@tue.nl

[1] Orthopaedic Biomechanics, Department of Biomedical Engineering, Eindhoven University of Technology, De Zaale, Groene Loper 15, 5612AP Eindhoven, The Netherlands

[2] Department of Mechanics and Production, Autonomous University of Manizales, Antigua Estación del Ferrocarril, Manizales, Caldas, Colombia

[3] Institute for Surgical Technology and Biomechanics, University of Bern, Stauffacherstrasse 78, 3014 Bern, Switzerland

Instead, an empirical failure criterion is used that assumes that failure will occur if the strain exceeds a critical value (typically set at 0.7%) for a prescribed amount of bone tissue (typically set at 2% of the tissue volume) [5]. Several validation studies have been performed to investigate the accuracy and precision of these predictions. In these validation studies, micro-FE based on HR-pQCT images of cadaver bones were used and the predicted results were compared to measured stiffness and failure load values [5–8]. These studies have shown excellent correlations between predicted and measured failure load values (with R^2 up to 0.92) and high accuracy [7].

With the introduction of a second-generation HR-pQCT scanner, the image processing has changed. Due to the higher resolution of this scanner, reconstructions are made at a smaller voxel size (60.7 μm) and a Gauss filtering followed by thresholding is used to generate the segmented image. Thresholds are set at different values for the cortical and trabecular compartment [9]. These changes affect the results of the micro-FE analyses. In an in vivo study comparing micro-FE results based on the first generation HR-pQCT (XCT1) and the second-generation HR-pQCT (XCT2) images, good correlations were found ($R^2 > 0.9$), but a significant bias existed in the results, with XCT2 underestimating the failure load at the radius when using the common failure criterion [10]. This underestimation was due to the fact that the threshold used for the XCT1 images was deliberately set too low, as the segmented image originally was used only for the trabecular number calculation, resulting in an overestimation of the amount of bone when using this same image as the basis for micro-FE analysis. Therefore, in order to get good agreement with experimental results, elastic and strength parameters were tuned to compensate for the overestimation of the amount of bone tissue. As the thresholds used for the XCT2 images are higher, resulting in a better representation of the actual mineralized phase, these values require new tuning in order to predict accurate values. In an earlier study [11], we directly compared the results of micro-FE analyses based on XCT2 images with experimental results and found similar results as reported in other earlier studies with regard to the underestimation of the predicted failure load when using XCT2 versus XCT1 images [10]. In our earlier study, however, the tested region was 20 mm, which is not the same as the standard clinically measured region (10.2 mm) and comprises bone more distal and more proximal of the standard scan region. Because the empirical failure criterion is also dependent on the resolution and the size of model [8], no tweaking could be performed in a way that would be relevant for the clinical scan region in that earlier study.

Another issue with the introduction of the XCT2 has been the increase in solution times due to the increased size of the micro-FE models. Whereas the XCT1-based models typically take between 4 and 8 h to solve, the XCT2-based models typically are in the range of 12 to 24 h. In the earlier validation study mentioned above [11], we also introduced an alternative homogenized FE (hFE) approach that can speed up the calculations and in addition enables non-linear analyses. With this approach, the bone is modelled as a continuum characterized by its density and fabric (a measure of the orientation of trabeculae). In that study, we found that the precision and accuracy of the failure load prediction from such hFE models are as good as those of micro-FE (mFE) models. Two issues, however, prohibit the clinical use of this approach. First, as mentioned before, the tested region in that study was not the same as the standard clinically measured region. Although this should be less of an issue with this mechanistic rather than empirical failure load prediction, a proper analysis of accuracy and cpu-time would require models based on the clinical region. Second, unlike mFE analyses, which are highly standardized, hFE analyses involve a large number of parameters related to the meshing (element size, density, type), the homogenization (parameters used, distance used), fabric (fabric tensor used, normalization used), material model parameters (moduli, strength, plasticity, and damage parameters, local/non-local formulations), and analysis (load-steps, convergence criteria). Hence, results are reproducible only when a highly standardized implementation is available. In the earlier study, we used an in-house developed software for most of these calculations. As this software is not available to other users, this would require other users to redevelop these procedures and steps. The accuracy of the calculations then would need to be re-tested, and some of the parameters may need tuning in order to get accurate results.

The overall aim of this study, therefore, was to develop a well-standardized approach for mFE and hFE analyses that can provide accurate results for bone failure load predictions based on XCT2 images of the distal radius that can be used by other HR-pQCT users. Although the approach developed here is similar to the approach used in our earlier study [11], the use of software already available to HR-pQCT users (e.g., for calculating fabric tensors) required to make some modifications. Furthermore, a more elaborate homogenization scheme was adopted and some modifications were made to material constants in order to ease their physical interpretation. The specific goals of this study therefore were, first, to identify the elastic and failure parameters for mFE and hFE analyses for models based on XCT2 images when using this new standardized approach; second, to investigate the accuracy of the failure load prediction by mFE and hFE models that represent the clinical (10 mm) region using these parameters; and third, to compare the mFE and hFE performance in relation to the computational costs for a standardized implementation.

Methods

Material

HR-pQCT images (XtremeCT II, Scanco Medical AG, Brüttisellen, Switzerland) at 60.7 μm resolution for 22 distal radius segments of 20.4 mm in size (2 stacks of 168 slices) were available from an earlier study and a detailed description can be found there [11]. Briefly, the segments were obtained from 12 pairs of fresh frozen anatomic specimens of human forearms that were cut at 5 mm and at 25 mm proximal of the distal subchondral plate and included the standard 10-mm clinical region. The study was approved by the ethics committee of the Medical University of Vienna. The donors (5 females, mean age 82.4 years, and 7 males, mean age of 75.1 years) had no bone-related diseases and had voluntarily donated their bodies to the Center of Anatomy and Cell Biology of the Medical University of Vienna. Images were obtained using the standard clinical setting (68 kVp voltage, 1460 μA, 43 ms integration time).

Image processing

The images were processed using the standard clinical workflow. In a first step, periosteal contours were generated using an automatic contouring algorithm. In the following step, the cortical and trabecular compartments were automatically defined and a Gauss filtering (sigma = 0.8, support = 1 voxel) and thresholding operation (320 mgHa/cm^3 for trabecular bone and 450 mgHa/cm^3 for cortical bone) were applied to generate a segmented image.

In addition, the standard clinical region was selected based on the original scout views as the region located 9 mm proximal of the reference line placed at the distal subchondral plate. To analyze the clinically relevant region, the 20.4-mm segments were cropped to one stack of 168 slices (10.2 mm).

Mechanical test data

The mechanical test data was also obtained from the earlier study [11]. In summary, the 20.4-mm segments were tested in compression at a rate of 5 mm/min in a servo-hydraulic test machine in which the upper platen was fixed by a ball-bearing to optimize load transfer. Platens were sand-blasted to minimize in-plane deformation during compression. Forces were measured using a load cell and the displacement and rotations of the upper platen were captured using an optical system. From the force-displacement curves, the stiffness and failure load were calculated.

Micro-FE analysis

Micro-FE models were generated directly from the segmented images using a voxel conversion technique [12] for both the 20.4- and 10.2-mm segments. All elements were assigned linear elastic material properties with a Young's modulus of $E = 10$ GPa and a Poisson's ratio of $\nu = 0.3$. Boundary conditions simulated a compression test at a compressive strain of 1% in which the transversal displacement at the loaded surfaces was suppressed (as in the experiment). The stiffness of the segment was calculated as the reaction force over the applied displacement.

Three parameters were determined such that the root-mean-square error (RMSE) between predicted and measured parameters for the 20-mm segments was minimized. For the stiffness, this was done by first calculating the slope β_0 of the least square regression line using a standard regression equation:

$$\beta_0 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \quad (1)$$

with x_i the stiffness values calculated from the micro-FE models, y_i the experimentally measured stiffness values, and $n = 22$ the number of samples. The Young's modulus then was calculated as $10/\beta_0$ GPa.

The second was the "critical volume" v_{crit} and the third parameter the "critical strain" ε_{crit} as defined in the original criterion described in Pistoia et al. [5]. With the original criterion, the critical volume was set to $v_{crit} = 2\%$ and the critical strain to $\varepsilon_{crit} = 0.7\%$. To find the optimal values for these constants for models based on XCT2 images, the critical volume parameter was increased from 0.7% to 1.2% in steps of 0.1% and the critical volume parameter from 2% to 7% in steps of 1%. For each possible combination, the RMSE values were calculated and the combination of parameters that provided the lowest RMSE was selected.

Homogenized-FE analyses

The homogenization approach was based on earlier studies [11, 13] but differed in several details. A detailed overview of the steps used in this study is provided here.

In a first step, periosteal and endosteal contours were generated automatically using a standard procedure implemented in the scanner software [14]. A mask image, further referred to as the full mask, then was created from the periosteal contour that represents the complete volume of the bone within the periosteal contour. Similarly, a mask was created from the endosteal contour that delineates the trabecular compartment (trabecular mask). By subtracting both masks, a mask of the cortical compartment (cortical mask) is obtained.

In a second step, the full mask was downscaled in all directions by a factor of 28 to a voxel size of 1.7 mm, and for each larger voxel, a volume fraction was calculated based on the original number of mask voxels it contains over the total number of voxels in the larger voxel. Each larger voxel with a volume fraction that exceeds 1% was converted to an eight-node brick element (Fig. 1). The downscale factor of 28 was selected based on a mesh convergence study (Online Resource 1) and on the fact that 28 is a divisor of 168, which is the number of slices per stack generated by the scanner, such that the full sample can be meshed with elements of isotropic size.

In the third step, homogenized properties were assigned to the elements. The homogenization was done separately for the trabecular and cortical compartments. For the trabecular compartment, a spherical region with a fixed radius of 2 mm was defined around the centroid of each element. For each element, the density D then was evaluated from the original high-resolution gray-level images by averaging the density of all voxels within the sphere region. Based on this density, a bone volume fraction ρ_{trab} was calculated as $\rho_{trab} = D/1200$, where 1200 mgHa/cm^3 is considered the density of cortical bone (Fig. 2a). In case the sphere would have extended outside the trabecular compartment, the volume fraction was calculated only for its part within the compartment and this value was assigned to the element (Fig. 2b). In case the element was only partly within the compartment (Fig. 2c), the bone volume fraction again was calculated for the part of the sphere within the trabecular compartment and assigned only to that part of the element. To do so, also the fraction of the element that is within the trabecular compartment f_{trab} was calculated and stored. The same homogenization was also used to calculate an element fabric tensor. The fabric was calculated from the segmented image using the standard mean intercept length (MIL) instead of the mean surface length (MSL) tensor as in the earlier study [11], because the MIL tensor is also used for the standard morphological analyses. The MIL fabric tensor was normalized such that its eigenvalues m_i sum to 3.

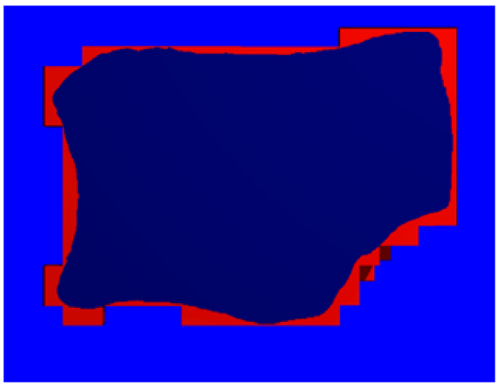


Fig. 1 The downscaled full mask (red) overlaid with the original full mask (dark blue). Note that the downscaled mask includes the full mask completely

Following, the homogenization was repeated for the cortical compartment. For this compartment, however, the density and fabric were averaged over the element volume rather than a larger spherical volume (Fig. 2d). In case the element volume was covered only partly by the cortical compartment, a cortical element fraction f_{cort} was calculated and stored (Fig. 2e, f). No fabric tensor was calculated for the cortical compartment. Instead, the fabric tensor was set to the identity tensor.

The final density of each element then was calculated as:

$$\rho = \frac{(f_{trab}\rho_{trab} + f_{cort}\rho_{cort})}{f_{trab} + f_{cort}} \tag{2}$$

Note that if an element is partly outside the periosteal contour, the element volume fractions f_{trab} and f_{cort} do not sum to 1, indicating the element is a "partial volume." As such, the density ρ represents the total density for that part of the element that is within the periosteal contour. A schematic representation of the procedure is shown in Fig. 3.

To find the averaged fabric tensor, a weighted arithmetic average was calculated of the fabric tensors calculated for the cortical and trabecular compartment [15]. With this procedure, the individual tensor components are averaged in the same way as described in Eq. 2. The averaged fabric tensor then was again normalized.

The material model used is the same as described in the previous paper. However, some of the parameters used in the earlier paper were changed to simplify their implementation and physical interpretation. Most notably, rather than a piecewise power function that relates the elastic parameters with volume fraction as well as the strength parameters with volume fraction, a single power function was used that also accounted for the element partial volume $f = f_{trab} + f_{cort}$. Accordingly, the three elastic moduli ε_i and three shear moduli μ_{ij} were calculated according to the relationship:

$$\begin{aligned} \varepsilon_i &= \varepsilon_0 f \rho^k m_i^{2l} \\ \mu_{ij} &= \mu_0 f \rho^k m_i^l m_j^l \end{aligned} \tag{3}$$

with ε_0 the elastic modulus of cortical bone, m_i the eigenvalues of the fabric tensor, the powers k and l material constants that were taken from the earlier study, and μ_0 the shear modulus of bone tissue. The tensile σ_i^+, compressive σ_i^-, and shear τ_{ij} strength calculations were adapted in a similar manner:

$$\begin{aligned} \sigma_i^{\pm} &= \sigma_0^{\pm} f \rho^p m_i^{2q} \\ \tau_{ij} &= \tau_0 f \rho^p m_i^q m_j^q \end{aligned} \tag{4}$$

with $\sigma_0^{\pm}$ the tensile/compressive strength and τ_0 the shear strength of bone tissue. Note that this formulation implies that moduli and strength values depend on the density and fabric according to a power function and on the partial volume according to a linear function.

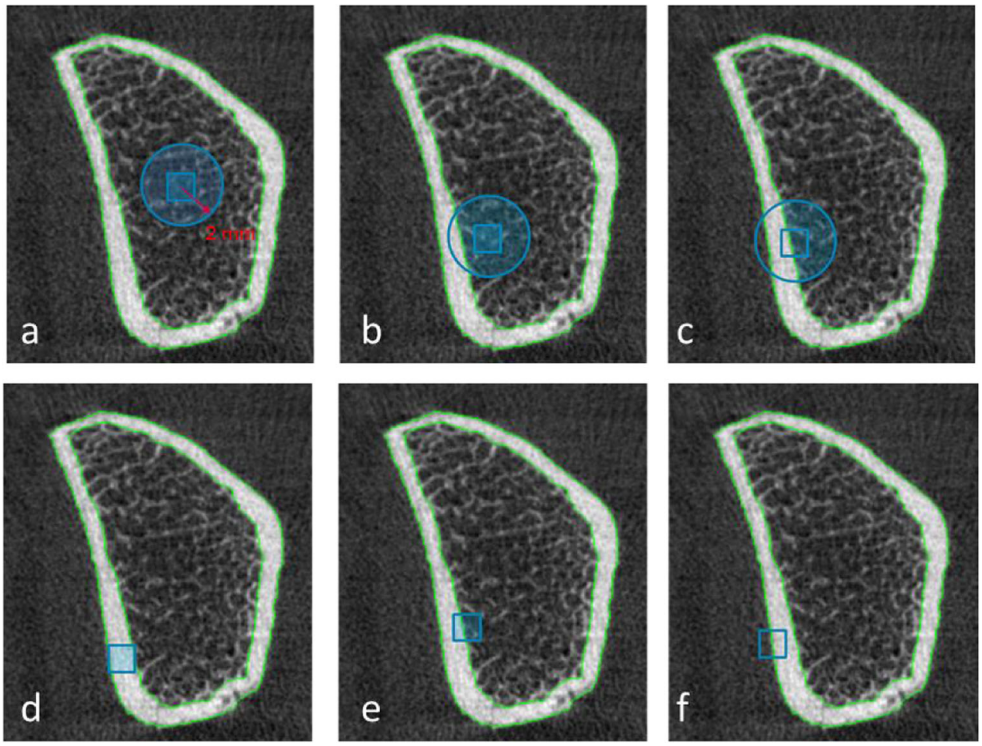


Fig. 2 Top row: Density homogenization of the cancellous region for a radius cross section with green lines delineating the periosteal and endosteal contours. The blue square represents a brick element and the blue circle the spherical region around the element centroid that is used for homogenization. Depending on the element location, different approaches are used: **a** If the spherical region was completely within the cancellous compartment, the element trabecular density (ρ_{trab}) was evaluated for the full sphere and assigned to the full element ($f_{trab} = 1$). **b** If the sphere was only partly inside the cancellous compartment, the element trabecular density (ρ_{trab}) was evaluated only for the part of the sphere within the cancellous region which was assigned to the full element ($f_{trab} = 1$). **c** If the sphere and the element were only partly within the cancellous compartment, the element trabecular density (ρ_{trab}) was evaluated only for the part of the sphere within the cancellous region which was assigned to the part of the element ($f_{trab} < 1$) that was within the cancellous compartment. Bottom row: Density homogenization of the cortical region. The blue square represents a brick element, which for the cortical bone also represents the region used for homogenization. Depending on the element location, different approaches were used: **d** If the element was completely within the cortical compartment, the element cortical density (ρ_{cort}) was evaluated for the full element region and assigned to the full element ($f_{cort} = 1$). **e, f** If the element was only partly within the cortical compartment, the element cortical density (ρ_{cort}) was evaluated only for the part of the element that was within the cortical compartment which was assigned to the part of the element ($f_{cort} < 1$) that was within the cortical compartment

As the material model differs in detail from the one used in the earlier publication, a validation study was performed first using the full 20-mm models for which mechanical test data was available. Using these models, two scaling factors were calculated to adjust elastic and strength parameters such that the RMSE between predicted and measured parameters for the 20-mm segments was minimized. The first scaling factor was for the two elastic constants (ε_0 and μ_0). The values used in the earlier study (but corrected for the absence of the piecewise power

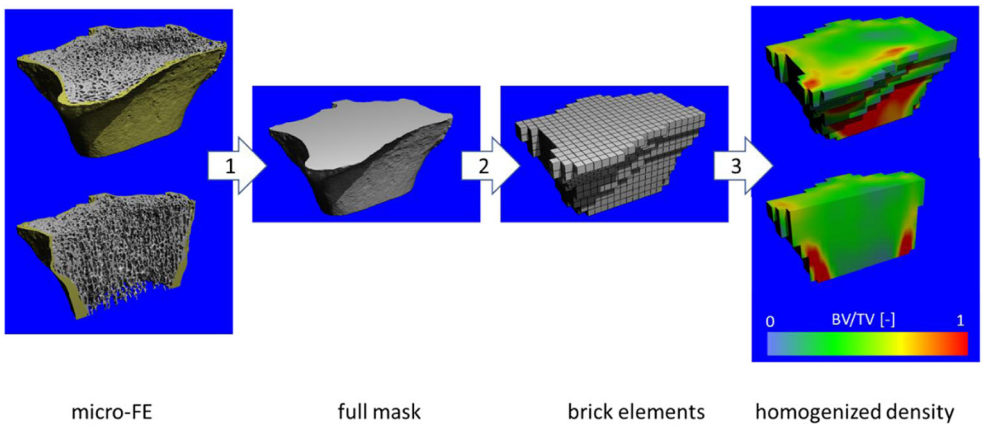


Fig. 3 Schematic representation of the homogenization procedure. In a first step, a "full mask" was generated representing the volume within the periosteal contour. In a second step, this mask was downscaled and segmented to obtain a 1.7-mm brick element representation. In a third step, the homogenized density and fabric were calculated for each brick element and assigned to the element

function) were used as a starting point and a single analysis then was performed for all models. The RMSE then was minimized by scaling the elastic constants with $1/\beta$, with β according to Eq. (1).

The second scaling factor was for the three strength constants (σ_0^-, σ_0^+, and τ_0). The values used in the earlier study were used as the starting point and a similar procedure was followed to calculate the scaling factor that would best match the calculated and measured failure load. Other parameters in the material model (i.e., the powers in the density relationships and the hardening constants) were taken from the previous study without tuning.

After fitting of the constants, the homogenized-FE analyses were repeated for the 10.2-mm sections and the predicted failure loads were compared to the experimentally measured ones. Since the stiffness largely depends on the height of the model, it was not considered useful to compare the stiffness values of the 10.2-mm sections to the measured stiffness values.

All image processing and finite element analyses steps were implemented in IPLFE v2.01 (Scanco Medical AG, Brüttisellen, Switzerland).

Statistics

Linear regression analyses were performed to compare the measured and calculated stiffness and failure load values:

$$y = \alpha + \beta x \tag{5}$$

with y the experimentally measured variable and x the FE-calculated variable, α the intercept, and β the slope of the regression line. A linear regression t test was used to test if the intercept was significantly different from zero and if the slope was significantly different from 1. Analyses were done using R version 3.4.2.

As 20 of the 22 samples were obtained from 10 donors (each providing a left and right sample), the samples cannot be considered fully independent. To check whether systematic differences exist between the left and right samples, a paired samples t test was performed.

Results

Parameter identification

For none of the parameters investigated, a significant difference between the experimentally measured and FE-calculated means left and right arm samples was detected, so the samples were further treated as independent.

Elastic parameters

For the mFE analyses, the RMSE for the predicted stiffness was minimized for a tissue Young's modulus of 10.83 GPa. Since the latter value was very close to the original value used in most earlier studies (10 GPa) and since the slope of the least square regression line when using this original value ($\beta_0 = 0.924$) was not significantly different from 1, it was decided to leave the tissue Young's modulus value to 10 GPa (Online Resource 2). The linear regression model then predicted the measured stiffness with a coefficient of determination $R^2 = 0.853$, a slope of 1.057, and an intercept that was not significantly different from zero (Fig. 4).

For the hFE analyses, the RMSE for the predicted stiffness was minimized when scaling the elastic constants by a factor of 1.534 (relative to the values reported in the earlier study that used a piecewise power function). The resulting moduli were 19.01 and 7.851 GPa for the Young's modulus and shear modulus, respectively (Online Resource 2). The linear regression model then predicted the measured stiffness with a coefficient of determination $R^2 = 0.856$, a slope that was not significantly different from 1, and an intercept that was not significantly different from zero (Fig. 4).

Strength parameters

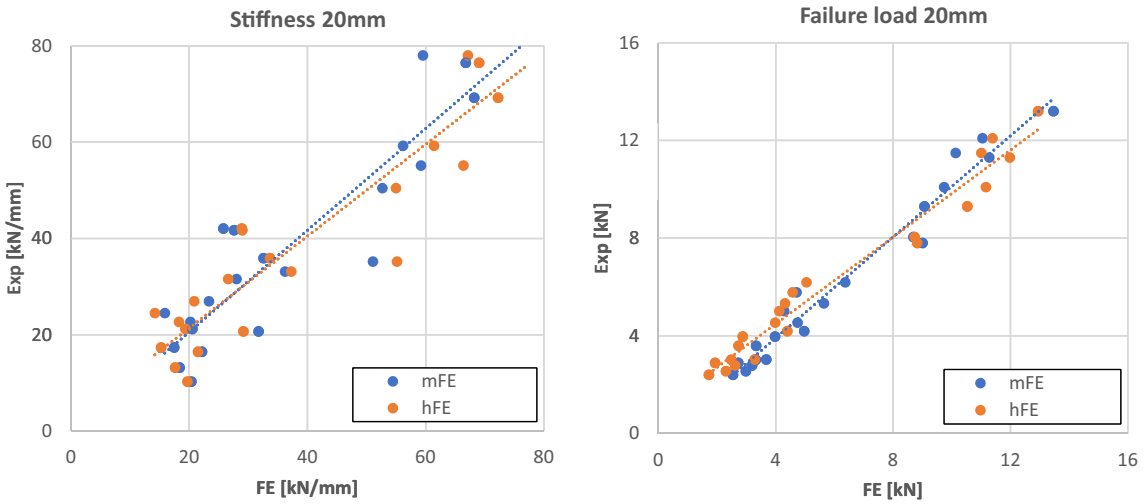
For the mFE analyses, the smallest RMSE was found when the critical strain was set to 1.0% and the critical volume to 5%. The final RMSE then was 620 N, which was similar to the best values found in an earlier study for XCT1-based analysis [8]. Using these parameters, the linear regression model predicted the measured failure load with a coefficient of determination $R^2 = 0.969$, a slope of 1.035, and an intercept that was not significantly different from zero (Fig. 4).

For the hFE analyses, using the moduli as described in the previous section, the RMSE for the predicted failure load was minimized when applying a scaling value of 0.753 for the strength parameters, resulting in a compressive strength $\sigma_0^- = 166$ MPa, a maximum tensile strength $\sigma_0^+ = 131$ MPa, and a maximum shear strength $\tau_0 = 67.3$ MPa (Online Resource 2).

Using these parameters, the linear regression model predicted the measured failure load with a coefficient of determination $R^2 = 0.966$, an intercept of 925 N, and a slope that was not significantly different from 1 (Fig. 4).

Failure load calculation for the clinical region

Using the set of identified constants (Online Resource 2), the failure load was calculated for the clinical region. For the mFE analyses, the linear regression model predicted the measured failure load with a coefficient of determination $R^2 = 0.955$, a



	Parameter	α	β	R^2
mFE	Stiffness	-588.8#	1.057	0.853
hFE	Stiffness	2299#	0.955*	0.856
mFE	Fult	-232.4#	1.035	0.969
hFE	Fult	925.0	0.890*	0.966

Not significantly different from 0, p<0.05
* Not significantly different from 1, p<0.05

Fig. 4 Regression of the FE-predicted stiffness (left) and failure load (right) with the experimentally measured values for the 20-mm segments

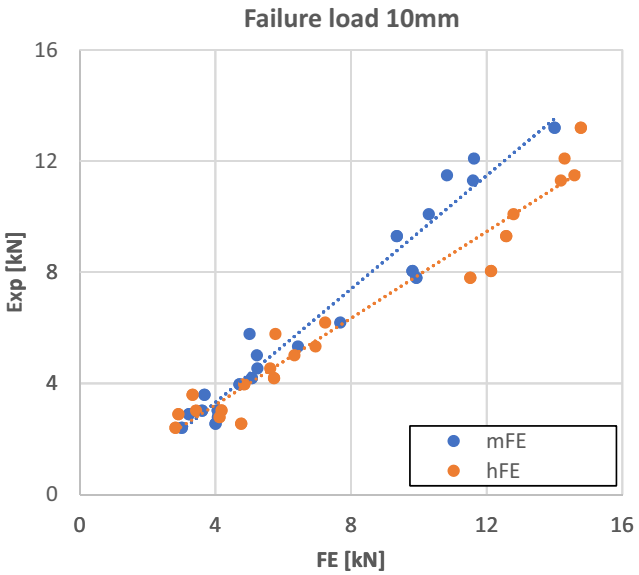
slope that was not significantly different from 1, and an intercept that was not significantly different from zero (Fig. 5).

For the hFE analyses, the linear regression model predicted the measured failure load with a coefficient of determination $R^2 = 0.952$, a slope of 0.780, and an intercept that was not significantly different from zero (Fig. 5).

A Bland-Altman plot was created to investigate differences between the failure loads predicted from the 20-mm segment models with those predicted from the clinical region models (Fig. 6, top). For all hFE models, an overprediction of the failure load was found (0.6–3.6 kN) that was higher for the stiffer segments. For the mFE models, a much smaller overprediction was found (0.06–0.75 kN) that remained constant over the range of failure loads measured. Two additional Bland-Altman plots were created to investigate differences between the failure loads predicted from both 20- and 10-mm segment models with those experimentally measured. For the 20-mm models, both mFE and hFE presented small differences (Fig. 6, bottom left), while for the 10-mm models, only the hFE models showed large differences (Fig. 6, bottom right).

Cpu-time results

All analyses were performed using 1 core of an HP Integrity Server rx2800. For the 10-mm segments, the average cpu-time required for solving the mFE models was 3 h and 45 m (SD = 1 h and 5 m). A full non-linear homogenized analysis with 1.7-mm elements took on average 0 h and 10 m (SD = 6 m), with roughly 30% of the



	Parameter	α	β	R^2
mFE	Fult	-770.6#	1.021*	0.955
hFE	Fult	102.4#	0.780	0.952

Not significantly different from 0, p<0.05
* Not significantly different from 1, p<0.01

Fig. 5 Regression of the FE-calculated failure load based on the 10-mm segments (horizontal axis) and the experimentally measured failure load for the 20-mm segments (vertical axis)

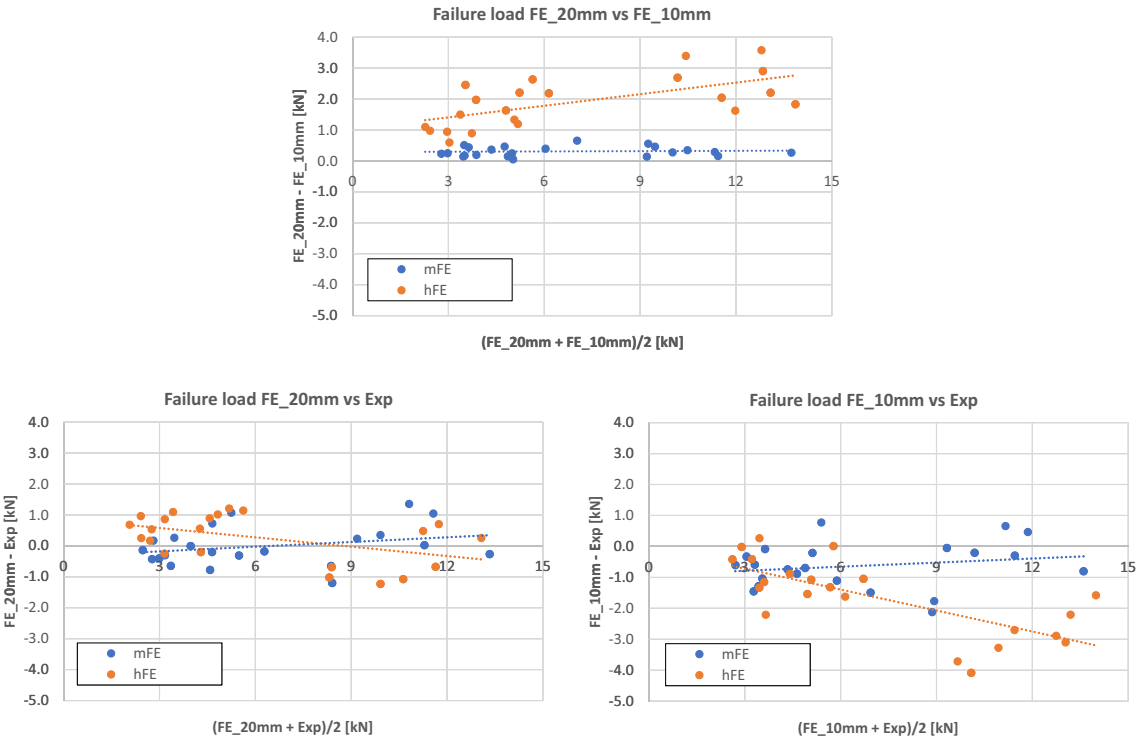


Fig. 6 Bland-Altman plots to compare the failure load calculated for the 20-mm segments with the failure load calculated for the 10-mm segments (top), the failure load calculated for the 20-mm segments with the experimentally measured failure load (bottom left), and the failure load calculated for the 10-mm segments with the experimentally measured failure load (bottom right)

cpu-time spent on solving of the hFE model and the rest on image pre-processing steps and homogenization. The cpu-time for solving the homogenized models, however, exponentially increased with reduced element size. Compared to the 1.7-mm elements, the cpu-time for solving the homogenized models with 1-mm models roughly increased by a factor 100 (Online Resource 1).

Discussion

The first goal of this study was to identify the elastic and failure parameters for mFE and hFE analyses when using models based on XCT2 images. For the mFE analyses, the default Young's modulus of 10 GPa provided accurate results. This value is in good agreement with values reported from bending, tension, and nanoindentation test of wet human trabecular bone [16] but is considerably lower than values typically reported for cortical bone (~ 19 GPa). It is also less than the value found when comparing the stiffness of elastic specimens obtained from micro-FE models with experimental measurement [16, 17]. Possible explanations for this discrepancy are the limited resolution of the models, the chosen threshold settings that will still result in some overestimation of the BV/TV [18], and the restrictive boundary condition used that can lead to an overestimation of the model stiffness (and hence underestimation of the tissue Young's modulus) [19]. As the Young's modulus used here was the same as that used in the earlier study [11], the results for the stiffness were the same as in that earlier study, although the regression constants are slightly different due to differences in statistical analysis. Our results are also in agreement with those of a recent study by Whittier et al. [20], who reported a tissue Young's modulus of 8748 MPa for XCT2 models. In their study, this value was obtained by comparing micro-FE results of models based on XCT2 images with those obtained from earlier validated XCT1 models of the same bones. The fact that their reported tissue modulus is still 13% less than the one we report potentially can relate to the fact that they did an in silico validation, whereas we performed an experimental validation. It is possible as well though that these differences relate to differences in the experimental setup used in the original validation experiment (in particular, the use of sand-blasted versus polished platens). Whereas no changes were needed for the tissue Young's modulus, the original mFE strength parameters according to Pistoia et al. [5] need to be changed as the original ones would lead to a severe underestimation of the bone failure load, a results that was reported as well by others [10, 20]. When using

the new constants as listed in Online Resource 2, an accurate estimation of the failure load was obtained with an excellent coefficient of determination ($R^2 = 0.9686$).

For the hFE analyses, results differ relative to the earlier study because of differences in implementation, homogenization approach, and fabric tensor used. As a result, the moduli and strength parameters identified in the present study differ substantially from those in the earlier study. The parameters found here, however, have a clearer interpretation as they directly represent the moduli/strength values of bone tissue. The values found for the tissue Young's modulus are in excellent agreement with the values reported for cortical bone (e.g., Mirzaali et al. [21] reported a value of 18.97 GPa) and with values reported for bone tissue in the literature obtained from nanoindentation (e.g., Rho et al. [22] reported a value of 19.4 GPa and Kim et al. [23] a value of 20 GPa for osteoporotic bone).

It should be noted that the failure load values measured in the earlier study and used in this study (average 6.3 kN) are somewhat higher than those reported in other studies [5, 7, 8]. These differences relate to differences in the material used and in the experimental setup. As a consequence, the failure loads predicted by the mFE and hFE using the parameters proposed here will be higher than those predicted in earlier normative studies [24, 25].

A second goal of this study was to investigate the accuracy of mFE and hFE models that represent the clinical 10-mm region using the new parameters. The results show that both models can predict the failure load of the 20-mm segments with a very similar and high coefficient of determination ($R^2 > 0.95$). The mFE models also provided good accuracy, with a non-significant intercept and a slope that was not significantly different from 1. The hFE models, on the other hand, tended to overpredict the failure load measured for the 20-mm segments when the model was restricted to the clinical section. A likely explanation for this overprediction is that in the experiment the actual failure occurred (partly) outside the clinical region. A plot of the accumulated damage (Online Resource 3) at the end of the loading curve revealed that, indeed, most of the damage for the 20-mm segments is expected at the distal boundary, while the predicted damage for the clinical segment is more uniform. This explanation agrees with earlier findings that demonstrated that more accurate predictions of distal radius failure load are possible if the clinical region is shifted further distally towards the region where many fractures occur [8]. The fact that this was not found for the mFE models might have several causes. First, the applied boundary conditions will stiffen the bone near the cut surfaces, which will reduce the deformations in particular in the weaker parts. Second, as these models are linear, the failure load is always based on the initial tissue strain distribution in the intact configuration. Third, the critical volume v_{crit}, in an absolute sense, is reduced in the clinical region models compared to the 20-mm segments, which can slightly increase the predicted failure load [8].

The third goal was to compare the mFE and hFE performance in relation to the computational costs. As the performance is very similar, only the cpu-time can be considered, which was around 22 times less for the hFE than for the mFE analyses. As presently only a standard FE solver is implemented, cpu-time for solving will exponentially increase when reducing the element size (Online Resource 1). Considering the results obtained in this study, however, there is no need to further reduce the element size.

A few limitations of the present study need to be discussed. First, 20 of the 22 test specimens were obtained from 10 donors and thus cannot be considered as fully independent samples. For that reason, the earlier study [11] calculated a marginal and a conditional coefficient of determination, which however complicates the comparison between the methods and with earlier studies. In the present study, instead, we checked if systematic differences could be detected between left and right arm parameters and further treated the samples as independent because we found no significant differences. Although the latter does not formally prove that the samples are independent, at least it provides some support for the assumption that no complications are to be expected by considering the samples as independent. An alternative option would be to make separate relationships for the left and right arm, but this was considered undesirable as there are no reasons to assume difference between left and right arm bone strength, which is also supported by the results of the *t* test. Averaging the left and right values would be statistically correct but obviously has the major disadvantage that it would halve the number of data points. Second, to simplify the workflow, calibration steps as performed in the previous study for BV/TV and the fabric tensor were not explicitly implemented in the present study but were implicitly accounted for by the tuning procedure. It is possible that accounting for the calibration would have changed the value of the constants, but the effects are expected to be very small. Third, the geometric representation of the models is coarse. The use of tetrahedron elements rather than the brick elements used here would have resulted in a much more accurate geometric representation. However, it was found that the results of such models using tetrahedron elements are almost identical to those of the present hexahedron (average difference in failure load was 0.7%, max. difference 4.3%) while the cpu-time for homogenization and solving is increased by approximately five times. As such, using such models for this application has no advantages other than these models provide a better geometric representation. Fourth, in the earlier study [11], the reproducibility of the hFE was found to be slightly less than that of the mFE approach. Although this was not tested, it is expected that the hFE

reproducibility would be improved in the present study because of the reduced number of processing steps and manual input. Finally, as demonstrated in the mesh convergence study, the failure load was not truly converged. Whereas this could be improved by using a non-local implementation, it is expected that this will not change the results much as the element size used here is close to the length scale that would be required for the non-local implementation.

As one of the overall aims of this study was to develop a well-standardized approach that can be used by others, the scripts and tasks to perform these homogenized- and micro-FE analyses will be made available as additional material (Online Resource 4) (Requires Scanco IPLFE v2.01 or higher to run).

In conclusion, by using the material parameters determined in this study, it is possible to accurately predict the stiffness and failure load using either a micro-FE or homogenized-FE approach. The latter has the advantage of being at least an order of magnitude faster.

Acknowledgements This study was supported by the grant program "Programa de Formación Doctoral Francisco José de Caldas Generación del Bicentenario" awarded by the Francisco José de Caldas Institute for the Development of Science and Technology (COLCIENCIAS, Colombia) LASPAUID20110290 and by funding (grant no. 14311.1 PFLS-LS) obtained from the Swiss Commission for Technology and Innovation CTI.

Compliance with ethical standards

Conflicts of interest Bert van Rietbergen is a consultant for Scanco Medical AG.

Andrés Julián Arias-Moreno, Hadi S. Hosseini, Melissa Bevers, Keita Ito, and Philippe Zysset declare that they have no conflict of interest.

Statement of human and animal rights For this type of study, formal consent is not required.

References

1. van Rietbergen B, Ito K (2015) A survey of micro-finite element analysis for clinical assessment of bone strength: the first decade. J Biomech 48(5):832–841
2. Samelson EJ, Broe KE, Xu H, Yang L, Boyd S, Biver E, Szulc P, Adachi J, Amin S, Atkinson E, Berger C, Burt L, Chapurlat R, Chevalley T, Ferrari S, Goltzman D, Hanley DA, Hannan MT, Khosla S, Liu CT, Lorentzon M, Mellstrom D, Merle B, Nethander M, Rizzoli R, Sornay-Rendu E, Van Rietbergen B, Sundh D, Wong AKO, Ohlsson C, Demissie S, Kiel DP, Bouxsein ML (2019) Cortical and trabecular bone microarchitecture as an independent predictor of incident fracture risk in older women and men in the Bone Microarchitecture International Consortium (BoMIC): a prospective study. Lancet Diabetes Endocrinol 7(1):34–43
3. Laib A, Rüegsegger P (1999) Comparison of structure extraction methods for in vivo trabecular bone measurements. Comput Med Imaging Graph 23(2):69–74
4. Boutroy S, van Rietbergen B, Sornay-Rendu E, Munoz F, Bouxsein ML, Delmas PD (2008) Finite element analysis based on in vivo HR-pQCT images of the distal radius is associated with wrist fracture in postmenopausal women. J Bone Miner Res 23(3):392–399
5. Pistoia W, van Rietbergen B, Lochmüller EM, Lill CA, Eckstein F, Rüegsegger P (2002) Estimation of distal radius failure load with micro-finite element analysis models based on three-dimensional peripheral quantitative computed tomography images. Bone 30(6):842–848
6. Macneil JA, Boyd SK (2008) Bone strength at the distal radius can be estimated from high-resolution peripheral quantitative computed tomography and the finite element method. Bone 42(6):1203–1213
7. Varga P, Pahr DH, Baumbach S, Zysset PK (2010) HR-pQCT based FE analysis of the most distal radius section provides an improved prediction of Colles' fracture load in vitro. Bone 47(5):982–988
8. Mueller TL, Christen D, Sandercott S, Boyd SK, van Rietbergen B, Eckstein F, Lochmüller EM, Müller R, van Lenthe GH (2011) Computational finite element bone mechanics accurately predicts mechanical competence in the human radius of an elderly population. Bone 48(6):1232–1238
9. Manske SL, Davison EM, Burt LA, Raymond DA, Boyd SK (2017) The estimation of second-generation HR-pQCT from first-generation HR-pQCT using in vivo cross-calibration. J Bone Miner Res 32(7):1514–1524
10. Agarwal S, Rosete F, Zhang C, McMahon DJ, Guo XE, Shane E, Nishiyama KK (2016) In vivo assessment of bone structure and estimated bone strength by first- and second-generation HR-pQCT. Osteoporos Int 27(10):2955–2966
11. Hosseini HS, Dünki A, Fabech J, Stauber M, Vilayphiou N, Pahr D, Pretterklieber M, Wandel J, van Rietbergen B, Zysset PK (2017) Fast estimation of Colles' fracture load of the distal section of the radius by homogenized finite element analysis based on HR-pQCT. Bone 97:65–75
12. van Rietbergen B, Weinans H, Huiskes R, Odgaard A (1995) A new method to determine trabecular bone elastic properties and loading using micromechanical finite-element models. J Biomech 28(1):69–81
13. Hazrati-Marangalou J, Ito K, van Rietbergen B (2015) A novel approach to estimate trabecular bone anisotropy from stress tensors. Biomech Model Mechanobiol 14(1):39–48
14. Burghardt AJ, Buie HR, Laib A, Majumdar S, Boyd SK (2010) Reproducibility of direct quantitative measures of cortical bone microarchitecture of the distal radius and tibia by HR-pQCT. Bone 47(3):519–528
15. Moahker M (2006) On the averaging of symmetric positive-definite tensors. J Elast 82:273–296
16. Wu D, Isaksson P, Ferguson SJ, Persson C (2018) Young's modulus of trabecular bone at the tissue level: a review. Acta Biomater 78:1–12
17. Bayraktar HH, Morgan EF, Niebur GL, Morris GE, Wong EK, Keaveny TM (2004) Comparison of the elastic and yield properties of human femoral trabecular and cortical bone tissue. J Biomech 37(1):27–35
18. Manske SL, Zhu Y, Sandino C, Boyd SK (2015) Human trabecular bone microarchitecture can be assessed independently of density with second generation HR-pQCT. Bone 79:213–221

19. Panyasantisuk J, Pahr DH, Zysset PK (2016) Effect of boundary conditions on yield properties of human femoral trabecular bone. Biomech Model Mechanobiol 15(5):1043–1053
20. Whittier DE, Manske SL, Kiel DP, Bouxsein M, Boyd SK (2018) Harmonizing finite element modelling for non-invasive strength estimation by high-resolution peripheral quantitative computed tomography. J Biomech 80:63–71
21. Mirzaali MJ, Schwiedrzik JJ, Thaiwichai S, Best JP, Michler J, Zysset PK, Wolfram U (2016) Mechanical properties of cortical bone and their relationships with age, gender, composition and microindentation properties in the elderly. Bone 93:196–211
22. Rho JY, Roy ME II, Tsui TY, Pharr GM (1999) Elastic properties of microstructural components of human bone tissue as measured by nanoindentation. J Biomed Mater Res 45(1):48–54
23. Kim G, Cole JH, Boskey AL, Baker SP, van der Meulen MC (2014) Reduced tissue-level stiffness and mineralization in osteoporotic cancellous bone. Calcif Tissue Int 95(2):125–131
24. Dalzell N, Kaptoge S, Morris N, Berthier A, Koller B, Braak L, van Rietbergen B, Reeve J (2009) Bone micro-architecture and determinants of strength in the radius and tibia: age-related changes in a population-based study of normal adults measured with high-resolution pQCT. Osteoporos Int 20(10):1683–1694
25. Burt LA, Schipilow JD, Boyd SK (2016) Competitive trampolining influences trabecular bone structure, bone size, and bone strength. J Sport Health Sci 5(4):469–475